Unilag Journal of Mathematics and Applications,
Volume 6, Issue 2 (2026), Pages 20--30.
ISSN: 2805 3966. URL: http://lagjma.edu.ng

A NEW APPROACH FOR MODELLING SKEWED-SEASONAL TIME SERIES DATASETS

OLUKUNMI OLATUNJI AKANNI*, TIMOTHY OLABISI OLATAYO AND ABASS ISHOLA TAIWO

ABSTRACT. New models applicable to skewed distributions have been developed since the transformation changes the structure of the original series. These models fail when applied to datasets that exhibit seasonality and skewness, as accurately modeling the data structure aids in forecasting and planning. The study proposed the Skewed-Seasonal Model (SSM) for the simulated data. The results showed that the proposed Skewed-Seasonal Model (SSM) accounted for the variability in the series better than the AR model. The Skewed-Seasonal Model (SSM) approach exhibited better goodness of fit, in addition to higher forecasting ability than the AR model. The forecast evaluation metrics indicated that the forecast evaluation of the Skewed-Seasonal Model (SSM) had lower values, making the proposed Skewed-Seasonal Model (SSM) more effective than the standard Autoregressive model in capturing and predicting the behavior of skewed-seasonal time series.

1. INTRODUCTION

Policymakers widely use time series analysis results as they aid in forecasting and decision-making [1,2]. High-dimensional data is still a rapidly developing field, particularly in the case of dependent data. Different models have been developed for time series data, but they are limited in application when the datasets possess certain characteristics [3,4,5]. Choudhury and Jones [6] examined several forecasting methods for assessing agricultural output estimates in Ghana to offer practical data for economic decision-making. They discovered that the ARMA model was superior to the other models and more robust (unaffected by cycle duration).

Fernandes and Salas [7] studied the gamma autoregressive (GAR(1)) model, which was one of the models that included the assumption that the underlying series had a gamma marginal distribution. Since the estimated parameters, obtained by means of the Method of Moments (MOM), are biased, some kind of correction was proposed before using them for parameter estimation. The authors suggested a technique for correcting the bias in the estimates of the model's parameters based on computer simulation studies. Adedotun et al. [8] proposed a model that can handle non-Gaussian

---

2010 *Mathematics Subject Classification*. Primary: 22E30. Secondary: 58J05.
Keywords and phrases. Skewed-Seasonal Model (SSM); Autoregressive Model; Durbin Watson; Autocorrelation function (ACF); Partial Autocorrelation function (PACF)

Submitted: January 20, 2026. Revised: February 14, 2026. Accepted: February 20, 2026.
*Correspondence

time series. The results indicated that NGAR(2) yielded a more accurate forecast. In the study on the Nigerian monthly exchange rate, the Non-Gaussian Autoregressive (NGAR) model clearly shows a very good alternative for analyzing non-normal time series data and, in particular, provides an efficient method of parameter estimation [9]. In a study on high-frequency data, Olatayo and Ekerikevwe [10] found that log-ARIMAX is more effective and outperforms the conventional ARIMAX model for observations that have heavy-tailed characteristics and outliers. Ribeiro et al. [11] proposed a bivariate generalized autoregressive (BGAR) model for non-normal data; the study showed that the new BGAR is more efficient than the competing models.

Harmonic analysis has been tried and tested in a number of geoscientific applications [12, 13, 14, 15] to estimate seasonal limits. In meteorology and climatology, harmonic analysis has also been extended [16, 17]. In order to explain seasonal trends in global surface air temperature, Legates and Willmott [18] used harmonic analysis. A study aimed at tackling issues in harmonic time series analysis for seasonal land surface dynamics by utilizing remote sensing data pointed out the challenges encountered in the model, such as the complexity of frequency selection, susceptibility to noise, and missing data in time series due to atmospheric interference [19,20,21,22].

Since the transformation changes the structure of the original series, new models suitable for skewed data have been developed [7, 8, 9, 10]. Modeling the seasonality of time series data is useful for understanding the effects of natural disturbances, as harmonic time series remain a suitable method. To capture the skewness and seasonality, there is a need to develop a robust model that accounts for the presence of these attributes in our dataset.

2. MATERIALS AND METHODS

The proposed model is given as:

$$y_t = \phi y_{t-1} + \beta \sin(\omega t) + \gamma \cos(\omega t) + \varepsilon_t \tag{2.1}$$

y_t represents a univariate variable with a constant interval of time observations.
β and γ represent the coefficients for the sinusoidal components
ω represents different frequencies for each harmonic component
ϵ_t represents the error term, $\varepsilon_{t\sim}N(0,\sigma^2)$

The model steps involve identification using the time plot and estimation of the coefficients using Maximum Likelihood Estimation and Ordinary Least Squares methods.

$$f(Y_t|Y_{t-1}, \sin(\omega t), \cos(\omega t), \phi, \beta, \gamma, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_t - \phi Y_{t-1} - \beta \sin(\omega t) - \gamma \cos(\omega t))^2}{2\sigma^2}\right) \tag{2.2}$$

$$L(\phi, \beta, \gamma, \sigma^2) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_t - \phi Y_{t-1} - \beta \sin(\omega t) - \gamma \cos(\omega t))^2}{2\sigma^2}\right) \tag{2.3}$$

$$\ell(\phi, \beta, \gamma, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{n} (Y_t - \phi Y_{t-1} - \beta \sin(\omega t) - \gamma \cos(\omega t))^2 \tag{2.4}$$

The estimates are:

$$\hat{\phi} = \frac{\sum_{t=1}^{n} \left(Y_t - \hat{\beta}\sin(\omega t) - \hat{\gamma}\cos(\omega t)\right)Y_{t-1}}{\sum_{t=1}^{n} Y_{t-1}^2} \tag{2.5}$$

$$\hat{\beta} = \frac{\sum_{t=1}^{n} \left(Y_t - \hat{\phi}Y_{t-1} - \hat{\gamma}\cos(\omega t)\right)\sin(\omega t)}{\sum_{t=1}^{n} \sin^2(\omega t)} \tag{2.6}$$

$$\hat{\gamma} = \frac{\sum_{t=1}^{n} \left(Y_t - \hat{\phi}Y_{t-1} - \hat{\beta}\sin(\omega t)\right)\cos(\omega t)}{\sum_{t=1}^{n} \cos^2(\omega t)} \tag{2.7}$$

Using the Ordinary Least Square method, the estimate is given in matrix form as

$$\begin{bmatrix} \sum Y_{t-1}^2 & \sum Y_{t-1}\sin(\omega t) & \sum Y_{t-1}\cos(\omega t) \\ \sum Y_{t-1}\sin(\omega t) & \sum \sin^2(\omega t) & \sum \sin(\omega t)\cos(\omega t) \\ \sum Y_{t-1}\cos(\omega t) & \sum \sin(\omega t)\cos(\omega t) & \sum \cos^2(\omega t) \end{bmatrix} \begin{bmatrix} \hat{\phi} \\ \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} \sum Y_t Y_{t-1} \\ \sum Y_t \sin(\omega t) \\ \sum Y_t \cos(\omega t) \end{bmatrix} \tag{2.8}$$

The optimal model criterion is given as [23]:

The AIC statistics is given as:

$$\text{AIC} = -2\ln \text{Likelihood}\left(\hat{\varphi}, \hat{\theta}, \hat{\sigma}^2\right) + 2(p + q + 1) \tag{2.9}$$

Bayesian information criteria (BIC) [24] is given as;

$$\text{BIC} = (n - p - q)\ln\left(\frac{n\hat{\sigma}^2}{n-p-q}\right) + n\left(1 + \ln\sqrt{2\pi}\right) + (p + q)\ln\left[\frac{\sum_{t=1}^{n} (X_t^2 - n\hat{\sigma}^2)}{p+q}\right] \tag{2.10}$$

The test statistic for sample autocorrelation is given as [25]:

$$Q(r) = n'(n' + 2)\sum_{j=1}^{k} \frac{\mu_j^2}{n-j} \tag{2.11}$$

The forecast accuracy measures are as follows [26]. The MAE is provided as:

$$\frac{1}{h+1}\sum_{t=s}^{h+s} \left(\hat{X}_t + X_t\right)^2 \tag{2.12}$$

RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{h+1}\sum_{t=s}^{h+s} \left(\hat{X}_t - X_t\right)^2} \tag{2.13}$$

and MAPE is given as

$$\text{MAPE} = \frac{100}{h+s}\sum_{t=s}^{h+s}\left|\frac{\hat{X}_t - X_t}{X_t}\right| \tag{2.14}$$

3. RESULT

The time plot showed the non-stationarity of the series with the ADF value (ADF = 10.265, p-value = 0.06). Furthermore, Figure 2 highlighted that the series demonstrated a skewed property. This suggests that the distribution of the observations is asymmetric. Given these characteristics, the proposed models in this study are deemed suitable for capturing the behavior of skewed seasonal observations (Figure 1). Figures 3 and 4 displayed the ACF and PACF of the series, respectively. The ACF plot exhibits a pattern where autocorrelations decay or decline exponentially, with a sharp cut-off in the PACF at lag k, suggesting the usage of an autoregressive model.

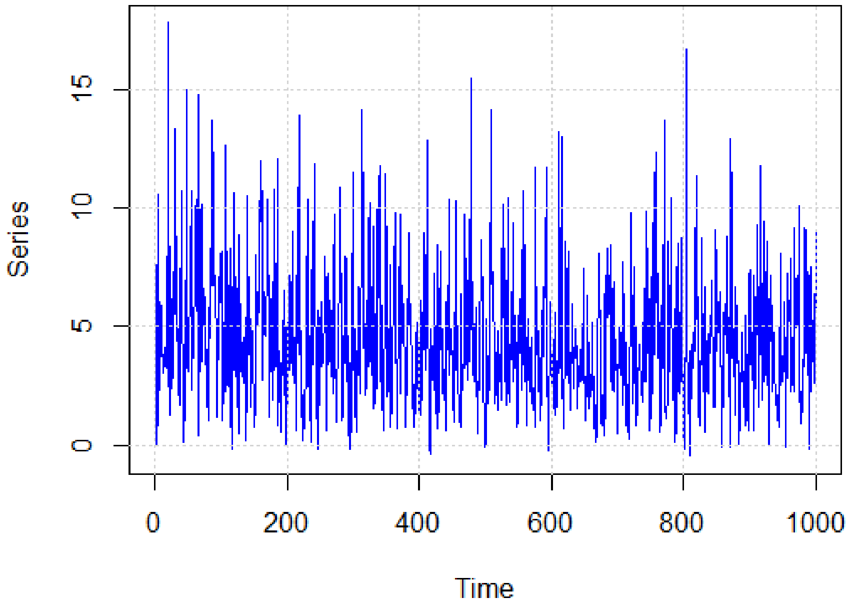

FIGURE 1: Time plot of the Series

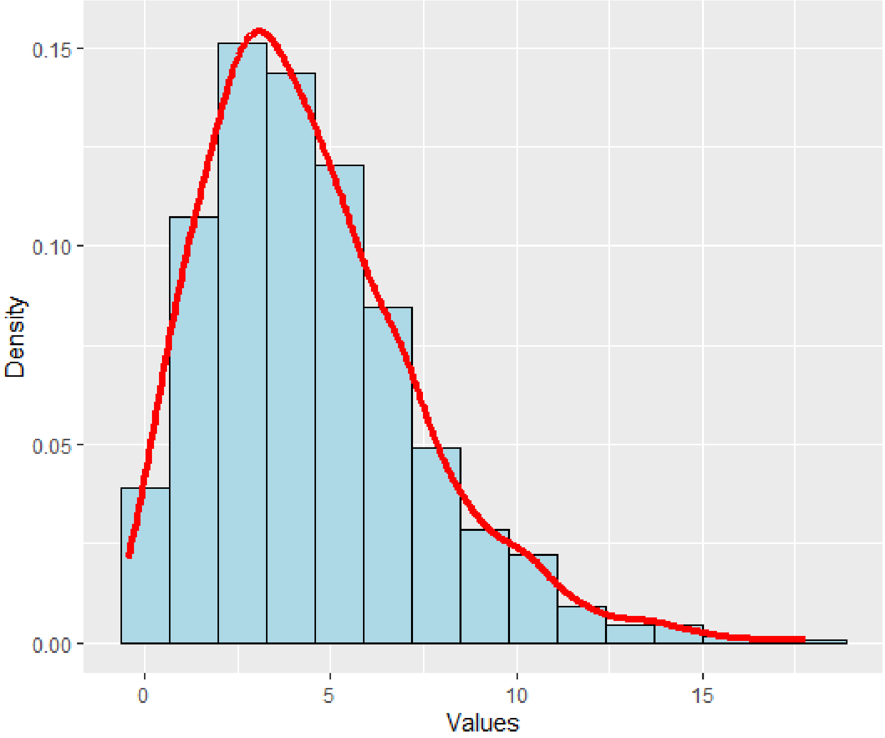

FIGURE 2: Histogram of the Series

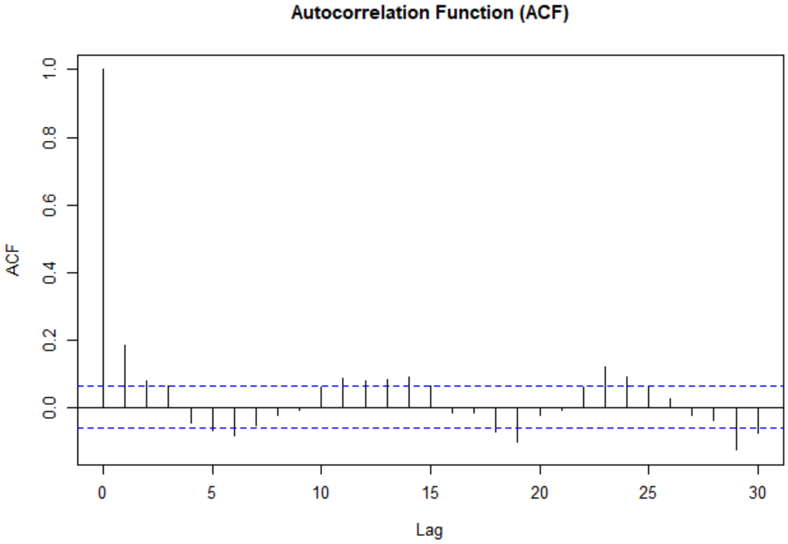

FIGURE 3: ACF of the Series

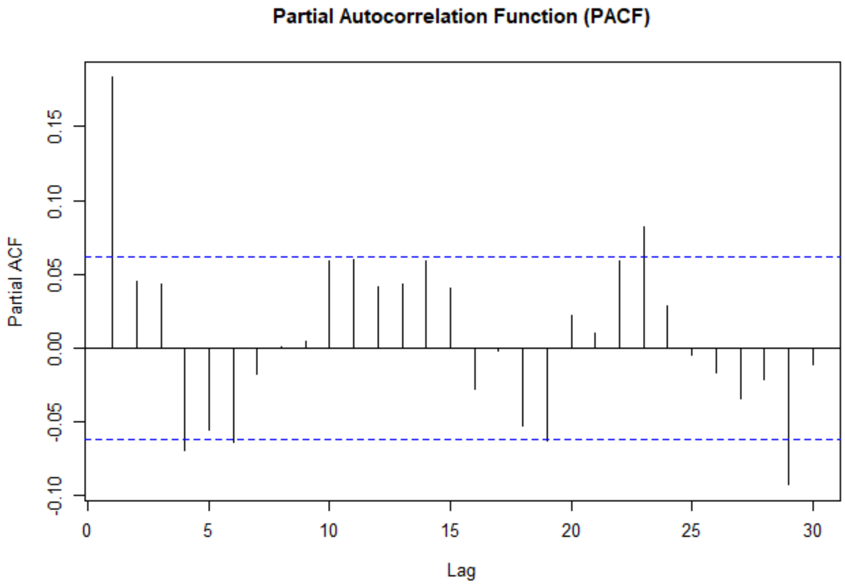

FIGURE 4: PACF of the Series

The Skewed-Seasonal Model (SSM) for order 1 estimated for the series is given as;

$$y_t = 4.4611 + \ 0.1188 y_{t-1} + 0.6602 sin \frac{2\pi_1}{12} t \ + \ 0.9991 cos \frac{2\pi_1}{12} t \ + \varepsilon_t \tag{3.1}$$

with $R^2 = 0.7872, Adjusted\ R^2 = 0.7591$, Durbin Watson $= 1.9991$, AIC $= \ 4877.7070$

The Skewed-Seasonal Model (SSM) for order 2 estimated for the series is given as;

$$y_t = 4.4613 + \ 0.1160 y_{t-1} + 0.02323 y_{t-2} + 0.6604 sin \frac{2\pi_1}{12} t \ + 0.01303 sin \frac{2\pi_2}{12} t$$
$$+ 0.9993 cos \frac{2\pi_1}{12} t \ + 0.007179 cos \frac{2\pi_2}{12} t \ \ + \varepsilon_t \tag{3.2}$$

With $R^2 = 0.8926, Adjusted\ R^2 = 0.8372$, Durbin Watson $= 1.9966$, AIC $= 4875.852$

The Skewed-Seasonal Model (SSM) for order 3 estimated for the series is given as;

$$y_t = 4.4608 + \ 0.1144 + 0.01773 y_{t-2} + 0.06304 y_{t-3} + 0.6603 sin \frac{2\pi_1}{12} t$$
$$+ 0.01432 sin \frac{2\pi_2}{12} t - 0.1755 sin \frac{2\pi_3}{12} t + 0.9987 cos \frac{2\pi_1}{12} t$$
$$+ \ 0.008270 cos \frac{2\pi_2}{12} t + 0.03074 cos \frac{2\pi_3}{12} t + \varepsilon_t \tag{3.3}$$

with $R^2 = 0.9472, Adjusted\ R^2 = 0.9131$, Durbin Watson $= 1.9913$, AIC $= 4875.684$

The Skewed-Seasonal Model (SSM) for order 4 estimated for the series is given as;

$$\begin{aligned} y_t = 4.4606 + {} & 0.1179 y_{t-1} + 0.01929 y_{t-2} + 0.06183 y_{t-3} - 0.01894 y_{t-4} \\ & + 0.6604 sin \frac{2\pi_1}{12} t + 0.01404 sin \frac{2\pi_2}{12} t - 0.1759 sin \frac{2\pi_3}{12} t - 0.04371 sin \frac{2\pi_4}{12} t \\ & + 0.9982 cos \frac{2\pi_1}{12} t + 0.007724 cos \frac{2\pi_2}{12} t + 0.03046 cos \frac{2\pi_3}{12} t \\ & - 0.2269 cos \frac{2\pi_4}{12} t + \varepsilon_t \end{aligned} \tag{3.4}$$

with $R^2 = 0.9832, Adjusted\ R^2 = 0.9742, \text{Durbin Watson} = 1.993737, \text{AIC} = 4870.401$

4. DISCUSSION

The Skewed-Seasonal Model (1) analysis showed that the coefficients indicated that for every increase in time, there would be a unit increase or rise in the series. The coefficient of determination in equation (3.1) revealed that time explained the variations in the simulated series by 78.7%. Additionally, the model's value showed a good goodness of fit with a 75.9% level of predictive power. The coefficient of determination in equation (3.2) revealed that time explained the variations in the simulated series by as much as 89.2%. Additionally, the model's value showed that it has a goodness of fit with an 83.7% level of predictive power. The coefficient of determination in equation (3.3) revealed that time explained the variations in the simulated series by as much as 94.7%. Additionally, the model's value showed that it has a good goodness of fit with a 91.3% level of predictive power. The coefficient of determination in equation (3.4) revealed that time explained the variations in the simulated series by as much as 98.3%. Additionally, the model value showed that it has a good goodness of fit with a 97.4% level of predictive power. The values of the Durbin-Watson statistics stated in equations (3.1-3.4) show that the error terms are not serially correlated, as the proposed Skewed-Seasonal Model (SSM) is considered adequate for the series.

From Table 1, the values for the SSM indicate that it explains more variance in the data compared to the AR model at all orders. Similarly, the adjusted values confirm that the Skewed-Seasonal Model (SSM) provides a better fit. The findings suggest that the Skewed-Seasonal Model (SSM) has a stronger ability to capture the underlying skewness of the time series than the AR model. The higher values and adjusted values indicate that incorporating harmonic terms improves model performance. Furthermore, model forecasting accuracy was evaluated using forecast evaluation metrics. The results in Table 2 confirm that the SSM model outperforms the AR model in predictive accuracy metrics (RMSE, MASE, MAPE, and MAE). These results highlight the effectiveness of the Skewed-Seasonal Model (SSM) in capturing the skewness and seasonality properties of the series.

TABLE 1: Values of Coefficient and Adjusted Coefficient of determination for Autoregressive Model (AR) and The Skewed-Seasonal Model (SSM) from order 1-4

Model Order	**R^2(AR)**	**Adjusted R^2 (AR)**	**R^2 (SSM)**	**Adjusted R^2 (SSM)**
1	0.5337	0.5328	0.7872	0.7591
2	0.6357	0.6337	0.8926	0.8372
3	0.7375	0.7346	0.9472	0.9131
4	0.7423	0.7384	0.9832	0.9741

Researcher's Computation, 2025

TABLE 2: Forecast Evaluation Metrics for Autoregressive Model (AR) and The Skewed-Seasonal Model (SSM) from order 1-4

Model Order	**Variance (AR)**	**AIC (AR)**	**RMSE (AR)**	**MAE (AR)**	**MASE (AR)**	**Variance (SSM)**	**AIC (SSM)**	**RMSE (SSM)**	**MAE (SSM)**	**MASE (SSM)**
1	8.1097	4935.978	2.8463	2.2368	0.7902	7.5648	4877.707	2.7490	2.1525	0.7604
2	8.0935	4935.981	2.8435	2.2319	0.7885	7.5607	4875.852	2.7483	2.1502	0.7596
3	8.0783	4935.100	2.8408	2.2264	0.7865	7.5141	4875.684	2.7398	2.1421	0.7567
4	8.0381	4933.133	2.8337	2.2173	0.7833	7.4843	4870.401	2.7344	2.1373	0.7550

Researcher's Computation, 2025

5. CONCLUSION

The Skewed-Seasonal Model (SSM) was proposed and estimated for orders 1 through 4. R^2, adjusted R^2, and forecast evaluation metrics were used to compare the proposed model and AR. The results demonstrated that the inclusion of harmonic terms improved the model's ability to capture the underlying structure of the data. The proposed Skewed-Seasonal Model (SSM) is more effective as it possesses higher performance and forecasting accuracy. The findings suggest that incorporating harmonic terms enhances the model's ability to describe time-dependent patterns, making it a suitable approach for analyzing similar real-world datasets. The proposed model will help policymakers, stakeholders, and the government by providing a better method to analyze skewed-seasonal data, which will aid in better forecasting and policy implementation.

Acknowledgment. All authors whose research works were cited and listed in the reference.

Authors Contributions. Olukunmi Olatunji Akanni: Conceptualization, Manuscript drafting, Methodology, Analysis, Interpretation, Formatting, and Editing. Timothy Olabisi Olatayo: Methodology, Analysis, Interpretation, Formatting, and Editing. Abass Ishola Taiwo: Methodology, Analysis, Interpretation, Formatting, and Editing. All the authors read and approved the final copy of this manuscript.

Authors' Conflicts of interest. Nil

Funding Statement. There is no external funding for this research.

REFERENCES

[1] P. Malik, A. Dangi, A. Singh, T. Asst, A. Pratap, S. Parihar, U. Sharma, L. Mishra. An Analysis of Time Series Analysis and Forecasting Techniques. International Journal of Advanced Research in Computer and Communication Engineering; (2023). 9. https://doi.org/6.0415/IJARIIE-21608

[2] M. Kolambe. Forecasting the Future: A Comprehensive Review of Time Series Prediction Techniques. Journal of Electrical Systems; (2024). 20. 575-586. https://doi.org/10.52783/jes.1478.

[3] F. Jonas, A.D. Haderlein, A.N. Peterson, L. Burkitt, D.B. Mareels. Autoregressive models for biomedical signal processing; (2023). arXiv.Org, abs/2304.11070. https://doi.org/10.48550/arXiv.2304.11070

[4] Y. Liu, J. Wang, V. Leiva, A. Tapia, W. Tan, S. Liu. Robust autoregressive modeling and its diagnostic analytics with a COVID-19 related application. Journal of Applied Statistics; (2023). https://doi.org/10.1080/02664763.2023.2198178.

[5] P.C. Padhan. Application of ARIMA Model for Forecasting Agricultural Productivity in India. Journal of Agriculture and Social Science; 2012. 8(2), PP. 50-56. https://doi.org/11–017/AWB/2012/8–2–50–56.

[6] A. Choudhury, J. Jones. Crop Yield Prediction Using Time Series Models. Journal of Economics and Economic Education Research; (2014). 15(3), pp. 53.

[7] B. Fernandez, J. Salas. Gamma-autoregressive models for streamflow simulation. Journal of Hydraulic Engineering; 1990. 116, 1403–1414. https://doi.org/10.1061/(ASCE)0733-9429(1990)116:11(140)

[8] A.F. Adedotun, A.I. Taiwo, T.O. Olatayo. A Re-parametrised Autoregressive Model for Modelling Gross Domestic Product. ANNALS of Faculty Engineering Hunedoara-

International Journal of Engineering; (2020). 84-188 http://eprints.covenantuniversity.edu.ng/id/eprint/17549

[9] A.F. Adedotun, T.O. Olatayo, G.O. Odekina. On Non-Linear Non-Gaussian Autoregressive Model with Application to Daily Exchange Rate, Journal of Physics: Conference Series; (2022). https://doi.org/10.1088/1742-6596/2199/1/012031

[10] T.O. Olatayo, K. Ekerikevwe. Performance Measures for Evaluating the Accuracy of Time Series Hybrid Model Using High Frequency Data. Britain International of Exact Sciences (BIoEx); (2022). 4. 244-259. https://doi.org/10.33258/bioex.v4i3.760.

[11] T. F. Ribeiro, A. P. Alencar, F. M. Bayer, "Bivariate generalized autoregressive models for forecasting bivariate non-Gaussian times series", (2025). Available: https://arxiv.org/pdf/2507.14442

[12] S.K. Padhee, S. Dutta. Spatio-temporal reconstruction of MODIS NDVI by regional land surface phenology and harmonic analysis of time-series. GIScience & Remote Sensing; (2019). 56 (8), 1261–1288. https://doi.org/10.1080/15481603.2019.1646977

[13] Q. Tarawneh. Harmonic analysis of precipitation climatology in Saudi Arabia. Theoretical Applied Climatology; (2016). 124: 205-217. https://doi.org/10.1007/s00704-015-1408-z

[14] J.O. Fantola, O.O Akanni. Forecasting COVID-19 Confirmed Cases: Time Series Analysis Perspective. The International Journal of Science & Technology; (2021). 9(5), doi: 0.24940/theijst/2021/v9/i5/ST2105-015

[15] T.A. Lasisi, O.O Akanni. Time Series Modeling of Water Quality Parameters. Journal of Environmental and Earth Sciences. (2020). 10(9), 41-49, doi: 10.7176/JEES/10-9-07

[16] H. Van Loon, R.L. Jenne, K. Labitzke. Zonal Harmonic Standing Waves. Journal of Geophysical Research; 1973. 78: 4463-4471. https://doi.org/10.1029/JC078i021p04463

[17] F. Arabi, S. Shojaei, M. Zare, H.R.G. Malamiri. Investigating the capability of the Harmonic Analysis of Time Series (HANTS) algorithm in reconstructing time series images of daytime and nighttime land surface temperature from the MODIS sensor. Spat. Inf. Res. 32, 425–439 (2024). https://doi.org/10.1007/s41324-023-00569-3

[18] D.R. Legates, C.J. Willmott. Mean seasonal and spatial variability in gauge-corrected, global precipitation. International Journal of Climatology; (1990). 10(2): 111-127. http://dx.doi.org/10.1002/joc.3370100202

[19] Z. Qiang, Z. Zhu, G. Xian, C. Li. A novel regression method for harmonic analysis of time series. ISPRS Journal of Photogrammetry and Remote Sensing; (2022). 48-61. https://doi.org/10.1016/j.isprsjprs.2022.01.006

[20] T.A. Lasisi, O.O Akanni. Water Quality Parameters: Statistical Point of View, Journal of Natural Sciences Research. (2020). 10(9), 75-85, doi: 10.7176/JNSR/11-18-06

[21] D.O. Kofoworola, O.O Akanni, A.J. Solagbade. Knowledge and perceived benefits of telemedicine adoption and online medical consultation among healthcare professionals at Ade-Oyo Maternity Hospital, Ibadan, Oyo State. Electronic Journal of Medical and Educational Technologies. (2024). 17(1)

[22] S.A. Agboluaje, O.O. Akanni. M.U. Abdulrahman, M.G. Adelaja. Knowledge and Treatment Adherence Determinants among Tuberculosis Patients in Ibadan South West Local Government Area. Researchjournali's Journal of Public Health. (2026). 12(1). 1-19

[23] H. Akaike. Information theory and an extension of maximum likelihood principle, Second International Symposium on Information Theory, Akademia Kiado. (1973). pp. 267–281.

[24] H. Akaike. A Bayesian analysis of the minimum AIC procedure. Annal Institute of Mathematical Statistics; (1978). A(30), 9–14.

[25] G.M. Ljung, G.E.P. Box. On a Measure of Lack of Fit in Time Series Models. Biometrika; 1978. 65(2), 297–303. https://doi.org/10.2307/2335207.

[26] T.O. Olatayo, A.I. Taiwo. Modeling and Evaluating Performances with Neural Network Using Climate Time Series Data. Nigerian Journal of Mathematics and Applications; (2016). 25; 205-216

OLUKUNMI OLATUNJI AKANNI*
DEPARTMENT OF STATISTICS, UNIVERSITY OF IBADAN, IBADAN, OYO STATE, NIGERIA.
E-mail address: oo.akanni@ui.edu.ng

TIMOTHY OLABISI OLATAYO
DEPARTMENT OF STATISTICS, OLABISI ONABANJO UNIVERSITY, AGO-IWOYE, OGUN STATE, NIGERIA.
E-mail address: bisi.olatayo@oouagoiwoye.edu.ng

ABASS ISHOLA TAIWO
DEPARTMENT OF STATISTICS, OLABISI ONABANJO UNIVERSITY, AGO-IWOYE, OGUN STATE, NIGERIA.
E-mail address: taiwo.abass@oouagoiwoye.edu.ng